

## RESEARCH INTEREST

I'm interested in speech and language processing. In particular, I am enthusiastic about building efficient and robust speech recognition systems in low-resource and multilingual settings. I'm also interested in understanding the intricacies of human-machine interactions towards building intelligent multimodal dialogue systems. Presently, I'm focussed on building speech data resources for Zambian languages.

---

## EDUCATION

**MSc., Machine Intelligence**, Africa Institute of Mathematical Sciences (AIMS), Kigali, Rwanda Oct 2019 – Mar 2021

- Recipient of the Google and Meta (formerly Facebook) full scholarship to pursue the African Masters of Machine Intelligence programme at the Africa Institute of Mathematical Sciences (AIMS) in Kigali, Rwanda.
- Supervisor: Prof. Antonios Anastasopoulos
- Date of completion: **31/03/2021**
- **Comment:** completed (both course work and research project) just waiting for official results.

**MSc., Information Technology Management**, Binary University, Puchong, Malaysia Apr 2013 – Oct 2014

- Specialization: **Information Technology**

**BSc., Computer Science**, University of Zambia, Lusaka, Zambia Apr 2007 – Aug 2011

- Specialization: **Computer Systems Engineering**
- 

## TEACHING EXPERIENCE

**Lecturer**, Department of Computer Science, University of Zambia, Zambia May 2015 – **Present**

- **Undergraduate Computer Science courses:**

CSC 2702 - Databases and Information System, University of Zambia	2021/2022
CSC 3712 - Advanced Databases, University of Zambia	2021/2022
CSC 4792 - Introduction To Data Mining, University of Zambia	2021/2022

**Instructor**, CISCO Academy, University of Zambia, Zambia Sep 2016 – Jun 2019

- Instructor of Python Programming, IT Essentials and Cisco Certified Network Associate (CCNA) Level 1 and Level 2 CISCO courses.

**Teaching Assistant**, Department of Computer Science, University of Zambia, Zambia Feb 2012 – May 2015

- Teaching assistant for Introduction to Databases course in the Department of Computer Science under the Staff Development Fellowship.
- 

## PERSONAL DEVELOPMENT

- **Mathematics for Machine Learning Specialization** by Imperial College London on Coursera Nov 2019  
Certificate of completion
  - **Deep Learning Specialization** by DeepLearningAI on Coursera Aug 2018  
Certificate of completion
  - **Machine Learning** by Stanford University on Coursera Jul 2018  
Certificate of completion
- 

## PUBLICATIONS

- **C. Sikasote**, A. Anastasopoulos, *BembaSpeech: A Speech Recognition Corpus for the Bemba Language*, Proceedings of the Language and Resource Evaluation Conference (LREC), 2022.

- S. H. Chiwamba, J. Phiri, P. O.Y. Nkunika, **C. Sikasote**, M. M. Kabemba, M. N. Moonga, *Automated Fall Armyworm (Spodoptera frugiperda, J.E. Smith) Pheromone Trap Based on Machine Learning*, Journal of Computer Science 2019, 15 (12): 1759.1779

---

## SCHOLARSHIPS AND AWARDS

- **African Master of Machine Intelligence Scholarship**, Kigali, Rwanda Sep 2019
- **Special Research Fellowship**, University of Zambia, Lusaka, Zambia Sep 2019
- **Binary University Scholarship**, Binary University, Malaysia Apr 2013
- **Staff Development Fellowship**, University of Zambia, Lusaka, Zambia Feb 2012
- **Zambia Higher Education Scholarship**, University of Zambia, Lusaka, Zambia Apr 2007

---

## GRANTS

- **Lead Researcher/Co-PI**, Multimodal Datasets for the Bemba Language Project Apr 2021 - Aug 2022

---

## TECHNICAL SKILLS

- **Programming Languages:** Python
- **Database technologies:** MS SQL, MySQL
- **Operating System platforms:** Windows, Linux
- **Machine Learning libraries:** Scikit-learn, Pandas, NumPy, SciPy
- **Data Visualization tools:** Matplotlib, tSNE
- **Deep Learning frameworks:** Pytorch, Tensorflow
- **Speech Recognition Tools:** DeepSpeech, FairSeq

---

## OPEN-SOURCE DATASETS

- **Zambezi Voice** : A multilingual speech corpus for Zambian languages containing speech recognition data resources for three languages of Zambia. [ON-GOING]  
Link: <https://github.com/unza-speech-lab/zambezi-voice>
- **BembaSpeech** : ASR corpus consisting of 24 hrs of read speech for the Bemba language spoken by over 30% of the population in Zambia.  
Link: <https://github.com/csikasote/BembaSpeech>
- **Multimodal Dataset for the Bemba Language** : Large scale multimodal dataset for the Bemba language containing multi-turn image-grounded dialogues amounting to over 180 hours of speech with corresponding Bemba transcription and translation in English. The project was funded by *Lacuna Fund*.  
Link: <https://github.com/csikasote/igc-mu-cibemba>
- **Multilingual ASR Corpus for African Languages** : Multilingual speech corpus consisting of over 60 hours of read speech for 19 African languages. I contributed 2 hours of read speech for Bemba language.  
Link: <https://github.com/besacier/AMMICourse/tree/master/STUDENTS-RETURN>

---

## STUDENT SUPERVISION AND MENTORING

- **Stanly Mwape**, Undergraduate, University of Zambia 2021/2022
- **Kalinda Siaminwe**, Undergraduate, University of Zambia, 2021/2022
- **Bangiwe Zulu**, Undergraduate, University of Zambia 2021/2022
- **Godfrey Mkandawire**, Undergraduate, University of Zambia 2021/2022

---

## COMMUNITY SERVICE

- **Conference Volunteer**

International Conference on Learning Representations (ICLR)	Mar 2021
International Conference on Artificial Intelligence and Statistics (AISTATS)	Apr 2021